

## REMARKS

This application has been reviewed in light of the Office Action dated January 13, 2005. Claims 1-4, 6, 7, 9, 10, 12-15, 17, 18, 20, 21, and 23-33 are pending in this application. Claims 1, 12, and 23 have been amended to define more clearly what Applicant regards as his invention. Claims 1, 12, and 23-27 are independent.

Applicant notes with appreciation the indication that Claims 9, 10, 20, and 21 would be allowable if rewritten so as not to depend from a rejected claim, and with no change in scope. Claims 9, 10, 20, and 21 have not been so rewritten because, for the reasons given below, their base claims are believed to be allowable.

Claims 1, 12, and 23 were rejected under 35 U.S.C. § 112, second paragraph, as being indefinite. Claims 1, 12, and 23 have been carefully reviewed and amended as deemed necessary to ensure that they conform fully to the requirements of Section 112, second paragraph, with special attention to the points raised in paragraph 5 of the Office Action. It is believed that the rejection under Section 112, second paragraph, has been obviated, and its withdrawal is therefore respectfully requested.<sup>1</sup>

Claims 1-4, 6, 7, 12-15, 17, 18, and 23-33 were rejected under 35 U.S.C. § 103(a) as being unpatentable over U.S. Patent No. 5,583,762 (Shafer), in view of U.S. Patent No. 6,330,574 (Murashita), and further in view of U.S. Patent 6,105,022 (Takahashi et al.).

---

<sup>1</sup>/Indeed, as part of a telephone interview one of the Applicant's attorneys conducted with the Examiner on February 8, 2005, a proposed amendment to Claim 1, for overcoming the Section 112 rejection, was discussed. That proposed amendment has now been incorporated into each of Claims 1, 12, and 23. The Examiner stated during the telephone interview that he would take into consideration all of the arguments presented in the telephone interview and discuss them with his supervisor. However, to date Applicant has not received a reply from the Examiner.

Claim 1 is directed to a document type definition generating method for generating a document type definition of a structured document containing document elements of a plurality of document element types. Each one of the plurality of document element types has a document element name and each document element has a start tag and an end tag. A physical structure judging step includes judging a physical similarity between the document elements in the structured document, wherein the judging of the physical similarity is based on the physical position of the start tag of each document element in the structured document. A semantic structure judging step includes judging a semantic similarity between the document elements by comparing a character string form located between the start tag and the end tag of each of the document elements. A document type definition generating step includes judging a similarity of the document element tags based on the results obtained in the physical structure judging step and the semantic structure judging step, and generating the document type definition unifying the document element names of similar document elements. The document type definition generating step includes a redundancy removing step of, when the physical structure and the semantic structure of a plurality of document elements, having tags different in element name, are judged as being of the same document element type in the physical structure judging step and the semantic structure judging step, excluding one document element name from a document type definition generating object based on the judgment results obtained in the physical structure judging step and the semantic structure judging step.

Shafer, as understood by Applicant, relates to a method for generating a grammar for a collection of sample document records and to a process for reducing the number of rules of such grammars. Shafer merely matches a “start tag” with the

corresponding “end tag”, as described on column 3, lines 14-29.

The Office Action, citing column 3, lines 30-40 of Shafer, states that Shafer “demonstrates a *physical structure judging step* when it states the ‘*determination is made as to whether each such acquired grammar element is combined with a rule.*” (Emphasis in original.)

As recited in Claim 1, the physical structure judging step includes judging a physical similarity between the document elements, *based on the physical position of a start tag of each document element*. The cited portion of Shafer does not make a determination whether an acquired grammar element is combined with a rule based on the physical position of a start tag of each document element. Accordingly, Shafer is not seen to teach or suggest this feature of Claim 1.

Nothing has been found in Shafer that would teach or suggest a physical structure judging step of judging a physical similarity between document elements in a structured document, wherein the judging of the physical similarity is based on the physical position of a start tag of each document element in the structured document, as recited in Claim 1.

Moreover, the Office Action concedes that Shafer “does not explicitly disclose judging a semantic structure of each document element.” The Office Action states that “Murashita teaches a special code discriminating unit for determining whether inputted coded data is a special code showing inputting of coded data of a tag (compare to... *“judging a similarity of the tags based on judgment results of said physical structure judging step”*). (Emphasis in original.) The Office Action cites column 6, lines 55-67 of Murashita.

Murashita, as understood by Applicant, relates to a technique for compressing and decompressing data. Murashita discusses examining tag elements in a structured document to distinguish the structures of the respective tags.

The cited portion of Murashita (i.e., column 6, lines 55-67) does not disclose the semantic structure judging step recited in Claim 1, which judges a semantic similarity between document elements by *comparing a character string form* located between the start tag and end tag of each of the document elements. Murashita merely discusses decoding coded data following a special code on the basis of a decode table. Accordingly, Murashita is not understood to teach or suggest the semantic structure judging step of Claim 1.

That is, nothing has been found in Murashita that would teach or suggest judging a semantic similarity between document elements by comparing a character string form located between a start tag and an end tag of each of the document elements, as recited in Claim 1.

Moreover, the Office Action concedes that Shafer and Murashita fail to disclose “a judging step, wherein after the judging step occurs, excluding one document element name from a document type definition takes place.” The Office Action also states that Takahashi et al. provides a judging method based on both semantic information and physical location.

Takahashi et al., as understood by Applicant, relates to text cataloging and searching in a text control system utilizing a computer. A method is discussed for cataloging a structured text in a set of structured texts, each of which has a logical structure, and a method for searching such a set of structured texts for specific text content.

The Office Action states at page 4 that Takahashi et al. “provides a means for excluding one document element if it is found to be the same as that found within the structured index.” However, the method of Claim 1 excludes a *document element name*, not a document element, from a document type definition generating object. In this sense, Applicant respectfully submits that the Examiner is misinterpreting the language of Claim 1, because he equates Takahashi et al.’s excluding one *document element* if it is found to be the same as that found within the structured index, with the above-recited feature of Claim 1. Accordingly, Takahashi et al. is not understood to teach or suggest this feature of Claim 1.

Nothing has been found in Takahashi et al. that would teach or suggest the feature of, when the physical structure and the semantic structure of a plurality of document elements, having tags different in element name, are judged as being of the same document element type, excluding one document element name from a document type definition generating object based on the obtained judgment results, as recited in Claim 1.

Accordingly, Claim 1 is seen to be clearly allowable over Shafer, Murashita, and Takahashi et al., whether considered separately or in any permissible combination (if any).

Independent Claims 12, and 23-27 recite certain features which are similar to those discussed above with respect to Claim 1 and therefore are also believed to be patentable over the cited references for the reasons discussed above.

A review of the other art of record has failed to reveal anything which, in Applicants’ opinion, would remedy the deficiencies of the art discussed above, as references against the independent claims herein. Those claims are therefore believed

patentable over the art of record.

The other claims in this application are each dependent from one or another of the independent claims discussed above and are therefore believed patentable for the same reasons. Since each dependent claim is also deemed to define an additional aspect of the invention, however, the individual reconsideration of the patentability of each on its own merits is respectfully requested.

This Amendment After Final Action is believed clearly to place this application in condition for allowance and, therefore, its entry is believed proper under 37 C.F.R. § 1.116. Accordingly, entry of this Amendment After Final Action, as an earnest effort to advance prosecution and reduce the number of issues, is respectfully requested. Should the Examiner believe that issues remain outstanding, it is respectfully requested that the Examiner contact Applicants' undersigned attorney in an effort to resolve such issues and advance the case to issue.

In view of the foregoing amendments and remarks, Applicant respectfully requests favorable reconsideration and early passage to issue of the present application.

Applicant's undersigned attorney may be reached in our New York office by telephone at (212) 218-2100. All correspondence should continue to be directed to our below listed address.

Respectfully submitted,

A handwritten signature in dark ink, appearing to read "L. P. Diana", is written over a horizontal line.

Leonard P. Diana

Attorney for Applicant

Registration No.: 29,296

FITZPATRICK, CELLA, HARPER & SCINTO  
30 Rockefeller Plaza  
New York, New York 10112-3801  
Facsimile: (212) 218-2200

NY\_MAIN 494836v1